

RECEIVED

NOV 14 2002

TECH CENTER 1600/2900



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/877,606

DATE: 11/04/2002

TIME: 13:37:57

Input Set : A:\1676002US1.txt

Output Set: N:\CRF4\11042002\I877606.raw

4 <110> APPLICANT: Lu, Min
 5 Ji, Hong
 7 <120> TITLE OF INVENTION: Antigen for Developing Neutralizing Antibodies to Human
 Immunodeficiency
 8 Virus
 11 <130> FILE REFERENCE: 1676.002US1
 13 <140> CURRENT APPLICATION NUMBER: 09/877,606
 14 <141> CURRENT FILING DATE: 2001-06-08
 16 <150> PRIOR APPLICATION NUMBER: US 60/210,322
 17 <151> PRIOR FILING DATE: 2000-06-08
 19 <160> NUMBER OF SEQ ID NOS: 9
 21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 46
 25 <212> TYPE: PRT
 26 <213> ORGANISM: Artificial Sequence
 28 <220> FEATURE:
 29 <223> OTHER INFORMATION: The amino acid sequence of the COOH-terminal extension of
 gp41.
 31 <400> SEQUENCE: 1
 32 Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Gly Thr Lys Ile Lys
 33 1 5 10 15
 34 Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr His Ile
 35 20 25 30
 36 Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Arg
 37 35 40 45
 39 <210> SEQ ID NO: 2
 40 <211> LENGTH: 78
 41 <212> TYPE: PRT
 42 <213> ORGANISM: Artificial Sequence
 44 <220> FEATURE:
 45 <223> OTHER INFORMATION: Linear amino acid sequence of a helical wheel projection of
 C45-pII.
 47 <400> SEQUENCE: 2
 48 Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
 49 1 5 10 15
 50 Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
 51 20 25 30
 52 Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Lys Ile Lys
 53 35 40 45
 54 Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr His Ile
 55 50 55 60
 56 Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Arg
 57 65 70 75

ENTERED

59 <210> SEQ ID NO: 3
60 <211> LENGTH: 52

RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/877,606

TIME: 13:37:57

Input Set : A:\1676002US1.txt

Output Set: N:\CRF4\11042002\I877606.raw

RECEIVED
NOV 14 2002
TECH CENTER 1600/2900

```

61 <212> TYPE: PRT
62 <213> ORGANISM: Artificial Sequence
64 <220> FEATURE:
65 <223> OTHER INFORMATION: Amino acid sequence of C52.
67 <400> SEQUENCE: 3
68 Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
69 1 5 10 15
70 Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
71 20 25 30
72 Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
73 35 40 45
74 Trp Phe Asn Ile
75 50
77 <210> SEQ ID NO: 4
78 <211> LENGTH: 85
79 <212> TYPE: PRT
80 <213> ORGANISM: Artificial Sequence
82 <220> FEATURE:
83 <223> OTHER INFORMATION: The chimera of the C52 peptide and an isoleucine-zipper
trimer (GCN4-pII).
85 <400> SEQUENCE: 4
86 Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
87 1 5 10 15
88 Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
89 20 25 30
90 Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
91 35 40 45
92 Trp Phe Asn Ile Lys Ile Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile
93 50 55 60
94 Leu Ser Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys
95 65 70 75 80
96 Leu Ile Gly Glu Arg
97 85
99 <210> SEQ ID NO: 5
100 <211> LENGTH: 7
101 <212> TYPE: PRT
102 <213> ORGANISM: Artificial Sequence
104 <220> FEATURE:
105 <223> OTHER INFORMATION: Most common sequence of amino acids at positions 669-675 of
gp160 (M group
106 HIV-1 strain).
108 <400> SEQUENCE: 5
109 Leu Trp Asn Trp Phe Asp Ile
110 1 5
112 <210> SEQ ID NO: 6
113 <211> LENGTH: 6
114 <212> TYPE: PRT
115 <213> ORGANISM: Artificial Sequence
117 <220> FEATURE:
118 <223> OTHER INFORMATION: Amino acid sequence immediately NH2-terminal to SEQ ID NO:5.
120 <400> SEQUENCE: 6

```

RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/877,606

TIME: 13:37:57

Input Set : A:\1676002US1.txt

Output Set: N:\CRF4\11042002\I877606.raw

```

121 Glu Leu Asp Lys Trp Ala
122 1 5
124 <210> SEQ ID NO: 7
125 <211> LENGTH: 7
126 <212> TYPE: PRT
127 <213> ORGANISM: Artificial Sequence
129 <220> FEATURE:
130 <223> OTHER INFORMATION: Amino acids 669-675 of gp41.
132 <400> SEQUENCE: 7
133 Leu Trp Asn Trp Phe Asn Ile
134 1 5
136 <210> SEQ ID NO: 8
137 <211> LENGTH: 7
138 <212> TYPE: PRT
139 <213> ORGANISM: Artificial Sequence
141 <220> FEATURE:
142 <223> OTHER INFORMATION: Amino acids at positions 669-675 of gp160.
W--> 144 <221> NAME/KEY: SITE
145 <222> LOCATION: 1
146 <223> OTHER INFORMATION: Xaa = Leu, Ile, Val, or Ser.
W--> 148 <221> SITE
149 <222> LOCATION: 2
150 <223> OTHER INFORMATION: Xaa = tryptophan or no amino acid present due to the
presence of a stop
151 codon at this position in the reading frame.
W--> 153 <221> SITE
154 <222> LOCATION: 3
155 <223> OTHER INFORMATION: Xaa = Asn, Ser, Thr, or Asp or no amino acid present due to
the presence of
156 a stop codon at an earlier position in the reading frame.
W--> 158 <221> SITE
159 <222> LOCATION: 4
160 <223> OTHER INFORMATION: Xaa = Trp or no amino acid present due to the presence of a
stop codon at
161 this position or at an earlier position in the reading frame.
W--> 163 <221> SITE
164 <222> LOCATION: 5
165 <223> OTHER INFORMATION: Xaa = Phe, Leu, Tyr, or Ser or no amino acid present due to
the presence of
166 a stop codon at an earlier position in the reading frame.
W--> 168 <221> SITE
169 <222> LOCATION: (6)...(6)
170 <223> OTHER INFORMATION: Xaa = Asp, Ser, Asn, Gly, Thr, Glu, or no amino acid present
due to the
171 presence of a stop codon at this position or an earlier position in the
172 reading frame.
W--> 174 <221> SITE
175 <222> LOCATION: (7)...(7)
176 <223> OTHER INFORMATION: Xaa = Ile or Met or no amino acid present due to the
presence of a stop
177 codon at an earlier position in the reading frame.
W--> 180 <400> 8
W--> 181 Xaa Xaa Xaa Xaa Xaa Xaa Xaa

```

182 1 5
184 <210> SEQ ID NO: 9

RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/877,606

TIME: 13:37:57

Input Set : A:\1676002US1.txt

Output Set: N:\CRF4\11042002\I877606.raw

185 <211> LENGTH: 52

186 <212> TYPE: PRT

187 <213> ORGANISM: Artificial Sequence

189 <220> FEATURE:

190 <223> OTHER INFORMATION: Mutant (N656L) amino acid sequence of C52.

192 <400> SEQUENCE: 9

193 Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr

194 1 5 10 15

195 Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys

196 20 25 30

197 Leu Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn

198 35 40 45

199 Trp Phe Asn Ile

200 50

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/877,606

DATE: 11/04/2002
TIME: 13:37:58

Input Set : A:\1676002US1.txt
Output Set: N:\CRF4\11042002\I877606.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 1, 2, 3, 4, 5, 6, 7

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 7
Seq#:4; Line(s) 83
Seq#:5; Line(s) 105
Seq#:8; Line(s) 150,155,160,165,170,176

VERIFICATION SUMMARY

DATE: 11/04/2002

PATENT APPLICATION: US/09/877,606

TIME: 13:37:58

Input Set : A:\1676002US1.txt

Output Set: N:\CRF4\11042002\I877606.raw

L:144 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:148 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:153 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:158 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:163 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:168 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:174 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:180 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:8
L:181 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0